

# **The effect of GDP on a country's CO<sub>2</sub> emission**

Georgia Institute of Technology

Heecue Kim

ECON 3161 Econometric Analysis

Prof. Shatakshee Dhongde

## **Abstract**

Recently, the interest on the environment has been important. Countries now gather to solve those environmental problems, such as Kyoto Protocol and Paris Agreement. One of the biggest factors that causes the global environmental issues is the emission of carbon dioxide. This paper assesses the impact of a country's development to the increase of amount of CO<sub>2</sub> emission, with hypothesizing that there will be a positive relationship. Single and multiple regression models depict that there are some significant positive relationships between CO<sub>2</sub> emission and increase of GDP per capita, meaning that countries tend to emit more CO<sub>2</sub> as they seek development.

## I. Introduction

Nowadays, sustainable development became one of important keywords for the world economy. Since the industrial revolution, the world has experienced a huge development under the use of fossil fuels. However, the trend has been changed since many of the problems of using fossil fuels that produces CO<sub>2</sub>, have caused critical problems that can direct to the human survivals, such as climate change. Many countries started to understand the climate risks from the CO<sub>2</sub> emission and started to restrict it. Under United Nations Framework Convention on Climate Change (UNFCCC), countries have signed to cooperate to prevent the long-term risks, represented by both Kyoto Protocol and Paris Agreements. In fact, many of the leading countries started to restrict the CO<sub>2</sub> emission by law.

However, unlike those developed countries that can manage to decrease the CO<sub>2</sub> emission by either education or technologies, most of the developing countries argue that they still need to emit enormous amount of CO<sub>2</sub>, since those alternative energy sources are too expensive to use. Many of those countries argue that it is unfair to restrict their use of fossil fuels after all those developed countries have used up and achieved the rapid economic development. They also argue that it is the developed countries that are emitting the high percentage of the total emission, not the developing countries, so that it is not fair to prohibit the CO<sub>2</sub> emission from developing countries.

Energy is an important factor for a country to develop, and even though there are many renewable energies that are using now, the energy receiving from fossil fuels still take the high percentage of total energy production. This means that the more the country has a capability to produce more energy, the economy growth of the country will be higher. As fossil fuel remains the high proportion of most of the energy sources used in the Earth, it is easily assumed that those countries with wealthier status are emitting more CO<sub>2</sub> than poorer countries. This paper tries to uncover the actual relationship between the country's GDP level and the amount of CO<sub>2</sub> emission. It is hypothesized that the higher GDP country will emit more CO<sub>2</sub>, since it has more capability to do so, assuming the argument by most of those developing countries' argument is true.

## II. Literature Review

There were many of previous works that tried to analyze the relationship between the CO<sub>2</sub> emission and the economic development. There were many specific approaches to discover the relationship, resulting various results.

Holtz-Eakin and Selden (1995) have shown that those countries with GDP per capita lower than \$6,900 tend to emit more CO<sub>2</sub>, and that did direct to the economic growth – increase of GDP. However, at the same time, they discovered that there is a diminishing marginal propensity to emit CO<sub>2</sub> as the

economy develops in a country. With the discovery of high tendency to use more energy and emit CO<sub>2</sub> in middle-to-lower countries that experiences a rapid economic growth as more as they start using energy and emitting CO<sub>2</sub>.

Wang et al. (2011) uncovered the relationship between the CO<sub>2</sub> emission and GDP growth, concentrated on China, one of the biggest CO<sub>2</sub>-emitting countries, from 1995-2007. They did find out that the use of fossil energy and emission of CO<sub>2</sub> had high relationship with the development of China during the time and predicted that “the amount of CO<sub>2</sub> emissions in China will not decrease in a quite long period of time since its economic output will keep growing. They further suggested the Chinese government to regulate to prevent further increase of CO<sub>2</sub> emissions as the Chinese economy keeps on growing. This indicates that as the GDP and the economic output of a country grows, the more use of energy is required, and at current situation, still major energy source is from fossil fuels, there is high possibility for those countries to emit more CO<sub>2</sub>, thinking that China has received rapid economic development, but at the same time its CO<sub>2</sub> emission also increased drastically.

Similarly, Saidi and Hammami (2015) discovered the effect of the economic growth on energy use to be positive, arguing that “economic growth, CO<sub>2</sub> emission and energy consumption are complementary.” They also figured out that this applies to most of the countries in the world, including Europe, North Asia, and Latin America and Caribbean region. However, 58 countries seem to be quite a less number to represent the whole world, this paper will distinguish from their research by concentrating on having more data from many different countries that can represent the phenomenon of the whole world.

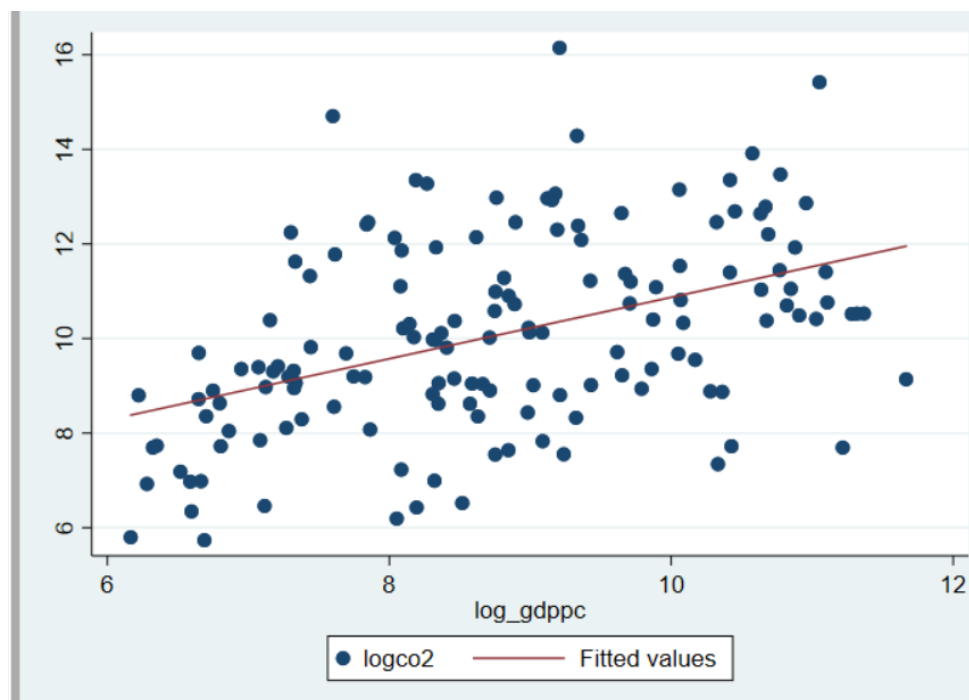
In this paper, it will concentrate on a specific year with cross-country data in order to figure out the actual relationship between countries. Most of those papers have proven the relationship through a time series data of a specific time period of a country or a region, and the effect of CO<sub>2</sub> produced for energy use impacting the developing economy. However, in this paper, it will be the other way round, looking for the relationship of countries having high GDP producing more CO<sub>2</sub>. To compare the actual emission difference between countries, the data from a specific, same time period is required.

### III. Data

To identify the relationship between economic growth and CO<sub>2</sub> emission of the whole world, cross-sectional data from a specific time-period is required. Data were gathered from the World Bank – World Development Indicators, where it provides data of 217 countries. However, there were several missing data from the file, the paper eliminated 69 countries that omitted their data. In order to gather as much recent data as possible, the year of 2018 was selected that had more available data compared to those of 2019 or after. The main dependent variable is the natural log of CO<sub>2</sub> emission to show the amount how much CO<sub>2</sub> each country has emitted. For the primary independent variable is the natural log of

GDP per capita, to measure the economic growth of a particular year. Data of 148 countries were gathered, and the list of countries are in the appendix.

Below is a scatter plot with a fit line of the natural log of CO2 emission and the GDP per capita, processed from STATA. The result depicts the general relationship between two, that one can figure out that the GDP increases, the CO2 emission tends to increase as well, indicating the positive relationship.



**Figure 1. Scatter plot between GDP per capita and CO2 emission**

For several other explanatory variables, this paper used renewable energy consumption, access to electricity, percentage of manufacturing factor from the GDP and urban population to create better regression model with multiple regression. The renewable energy consumption is chosen with straightforward reason that if there is more consumption of renewable energy, it will be high possibility for a country to use less energy from that emits CO2. There is also high possibility that low GDP countries having trouble receiving renewable energy source, as explained above, which can show some different results. For the access to electricity, it is thinkable that those country with less access to electricity is having lack of energy, which would not create much CO2 emission. For the percentage of manufacturing factor, it is reasonable to think that the country that has high percentage of manufacturing factor will have higher number of factories that will lead to the increase of CO2 emission. The number of urban populations is chosen because cities usually use more energy compared to those countryside areas. Therefore, it is assumed that the more people live in the urban area, there will be more use of energy creating higher CO2 emission. The tables below illustrate the further information about the variables and data that are used for this paper. Finally, for the further understanding of the difference

from the developed and developing countries, a dummy variable is further added depending on the country is the member of OECD or not. For the OECD dummy variable, the 5 key partner countries are also included. Appendix A elaborates more on each variables.

Variable Name	Description	Year	Units	Source
logco2	Natural log of CO2 emission	2018	kt	World Bank
log_gdppc	Natural log of GDP per capita	2018	Current USD	World Bank
renewable	Renewable energy consumption (% of final energy consumption)	2018	Percentage	World Bank
accelec	Access to electricity	2018	Percentage	World Bank
manuf	Manufacturing factor (% of GDP)	2018	Percentage	World Bank
log_urbanpop	Natural log of urban population	2018	People	World Bank
oecd	OECD member status	2021	0: not a member 1: a member	OECD

**Table 1. Variable description**

	# of Observations	Minimum	Maximum	Average	Standard Deviation
logco2	148	5.74	16.15	10.07	2.06
log_gdppc	148	6.17	11.67	8.77	1.41
renewable	148	0	96.38	32.75	27.74
accelec	148	10.12	100	85.15	24.94
manuf	148	1.69	39.91	12.82	6.47
log_urbanpop	148	12.07	20.53	15.65	1.67
oecd	148	0	1	0.27	0.45

**Table 2. Summary of statistics for the variables**

For the simple regression model between logco2 and log\_gdppc, and for the multiple regression model with more variables the Classical Linear Model (CLM) assumption should be verified. The assumptions are as below.

Assumption 1: Linear in parameters

The model will follow the assumption that is linear in parameters, as below.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

Assumption 2: Random Sampling

Since the data were collected from the World Bank, all of them are collected in random population and samples from the world, achieving the assumption.

Assumption 3: No perfect collinearity

Appendix B illustrates the result of correlation between the variables. It is found that there is no perfect collinearity that the value equals to 1. However, there were some high numbers that are getting closer to 1, so further determination will be needed in the robustness testing section.

Assumption 4: Zero conditional mean

It is true that there will be many variable factors that will affect the variables. However, for those multiple linear regression used in this paper, it is assumed to have zero conditional mean to the residuals – meaning that  $E[u|x_i] = 0$  for all  $i = 1, 2, \dots, n$ . Further determination can be done by the omitted variable bias  $\delta_i$  for all slopes from the variables. If the bias value is positive, then there will be a overestimation, vice versa.

Assumption 5: Homoskedasticity

Similar to the assumption 4, for the multiple linear regression, it is assumed that the expected variance of residual  $u$  is constant to any given dependent variables, meaning that  $V(u|x_i) = \sigma^2$  for all  $i = 1, 2, \dots, n$ . Appendix C illustrates the residual and fitted value graph for all variables. One can see that the spread of residual is equidistantly scattered from the zero line within the value of  $\pm 1$ . For those points that is over  $\pm 1$  need to be considered carefully.

Assumption 6: Normality of error terms

Finally, MLR 6 assumes that the  $u$  value is independent from explanatory variables and normally distributed. To determine for this assumption, histogram with the normal density curve and Q-Q plot is used to assist. The figures are provided in Appendix D, providing that the errors are forming

close to the normal distribution.

#### IV. Results

The first result is the simple regression model to identify the direct relationship between the CO2 emission and the GDP per capita without any further variables. This will depict the impact of GDP growth to the increase of CO2 emission. Each regression models will provide the equation and the standard error for each parameter inside the parentheses. Also, the n stands for number of observations and  $R^2$  as the sum of squared residuals. The further data, such as adjusted  $R^2$  are provided in appendix E-G below. All data were progressed under STATA.

##### **Simple Regression Model 1: $\log_{co2} = \beta_0 + \beta_1 \log\_gdppc + u$**

Regressing  $\log_{co2}$  on  $\log\_gdppc$ , the equation results as below.

$$\text{Equation 1: } \log_{co2} = 4.38 + 0.65 (\log\_gdppc)$$

$$(0.96) \quad (0.11)$$

$$n = 148 \quad R^2 = 0.20$$

This equation illustrates the relationship between CO2 emission and GDP per capita, visually shown above with the fit graph. As the fit graph has shown, there is a positive relationship between two, with 1 percent increase of GDP per capita leading to about 0.65 percent increase of CO2 emission. This depicts that the more a country's economy grows, that country emits more CO2 for the development. The R-squared value is 0.20, which means the regression explains 20% of the variation in CO2 emission. In addition, the t-value of  $\log\_gdppc$  is 5.99 with p-value of 0.00. This means that this regression model is statistically significant with below 1% level, which is a very promising level.

##### **Multiple Regression Model 2: $\log_{co2} = \beta_0 + \beta_1 \log\_gdppc + \beta_2 \text{renewable} + \beta_3 \text{accelec} + \beta_4 \text{manuf} + \beta_5 \log\_urbanpop + u$**

The first multiple regression equation is as below.

$$\text{Equation 2: } \log_{co2} = -8.62 + 0.32 (\log\_gdppc) - 0.014 (\text{renewable}) + 0.012 (\text{accelec}) +$$

$$(0.60) \quad (0.044) \quad (0.0024) \quad (0.0030)$$

$$0.0085 (\text{manuf}) + 0.97 (\log\_urbanpop)$$

$$(0.0072) \quad (0.027)$$

$$n = 148 \quad R^2 = 0.94$$

This equation depicts the relationship among CO2 emission, GDP per capita, renewable energy consumption, access to electricity, manufacturing factor percentage and urban population of a country. The R-squared value increased to 0.94, which means 94% of the dependent variable can be explained by variables of all explanatory variables of log\_gdppc, renewable, accelec, manu and log\_urbanpop. This significant increase of R-squared value can be highlighted that with those added explanatory variables. Unlike the relationship between GDP per capita, which is 0.32 (meaning that 1% increase leading to 0.32% increase of CO2 emission), the energy consumption of renewable energy has a negative relationship to the CO2 emission, that 1% increase of renewable energy will lead to 0.014% decrease of CO2. This data is interesting that it seems the increase of renewable energy causing less impact to the decrease of CO2 emission. This part can be studied further about the actual impact of renewable energy usage to CO2 emission. For the statistical significance of this model can be determined through the t-values and p-values of each variable. Especially highlighting on p-values, except for manu, which had 0.238 meaning that this is significant at the level of 23.8% (showing not quite insignificant), all other variables had 0.00 p-values, showing that everything else are providing promising statistical inferences.

### Multiple Regression Model 3:

$$\log_{co2} = \beta_0 + \beta_1 \log\_gdppc + \beta_2 \text{renewable} + \beta_3 \text{accelec} + \beta_4 \log\_urbanpop + u$$

The second multiple regression equation is as below.

Equation 3:

$$\log_{co2} = -8.70 + 0.32(\log\_gdppc) - 0.014(\text{renewable}) + 0.013(\text{accelec}) + 0.98(\log\_urbanpop)$$

(0.60)
(0.044)
(0.0024)
(0.0030)
(0.026)

n = 148
R<sup>2</sup> = 0.94

For this model, this paper eliminated manu variable that showed high p-value and regressed again. For this model, the R-squared value has weakly decrease by 0.0007, which means the elimination of manu did not impact the explanation of the model in critical level. Even though the rate has decreased, there is still a positive relationship with GDP per capita and negative relationship with renewable energy consumption. For this final multiple regression model, all t-values and p-values show promising numbers, provided in appendix G.



Overall, the table below summarizes those three models.

Dependent Variable: logco2			
Independent variables	Model 1	Model 2	Model 3
log_gdppc	0.65*** (0.11)	0.32*** (0.044)	0.32*** (0.044)
renewable		-0.14*** (0.0024)	-0.14*** (0.0024)
accelec		0.12*** (0.0030)	0.13*** (0.0030)
manuf		0.0085 (0.0072)	
log_urbanpop		0.97*** (0.027)	0.98*** (0.026)
Intercept	4.38*** (0.96)	-8.62*** (0.60)	-8.70*** (0.60)
Number of observations	148	148	148
R-squared	0.20	0.94	0.94
Adjusted R-squared	0.19	0.93	0.93

\*significant at \*10%, \*\*5%, \*\*\*1%

**Table 3. Estimation Results**

## V. Extensions

### A. Robustness test

As mentioned above in the assumption part, there were several parts that are doubted to be correlated. Therefore, *renewable* and *accelec* are chosen as they have provided the highest number in the table.

The null hypothesis for the robustness test will be as below.

$$H_0: \beta_2 = \beta_3 = 0$$

Model 2 was used for the unrestricted model, and below equation will be used as the restricted model with the two variables dropped. Those unrestricted and restricted models will be used for the F-test.

*restricted mode*:  $\log_{co2} = -11.36 + 0.63 \log\_gdppc + 0.016 \text{manuf} + 1.00 \log\_urbanpop$

The F statistic for the model follows.

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n - k - 1)} = \frac{(70.95 - 39.06)/2}{39.06/(148 - 5 - 1)} = 57.97$$

At the 1% significance level, the critical value for  $F_{2,142}$  is about 4.76. The value of 57.97 is way much higher than the F value and rejects the null hypothesis even in very small significance level. Therefore, these two variables are jointly significant, having multicollinearity. This should be carefully on account when one understands the model provided in this paper, with MLR 3 being ruined.

#### B. Different functional form

For further research of the precise impact of countries, it is better to see the regression model without the extreme values. Appendix H provides some further information about the scatter plot provided above in Figure 1. This paper figured out that the four major countries of United State, Russia, China, and India, which are mostly mentioned in the literature review, as the countries that are providing extreme values. Therefore, a regression model without those four countries is created and the result is as below.

#### Extra Multiple Regression Model 1

$$\log_{co2} = \beta_0 + \beta_1 \log\_gdppc + \beta_2 \text{renewable} + \beta_3 \text{accelec} + \beta_4 \log\_urbanpop + u$$

Equation 4:

$$\log_{co2} = -8.31 + 0.31(\log\_gdppc) - 0.013(\text{renewable}) + 0.013(\text{accelec}) + 0.95(\log\_urbanpop)$$

$$(0.62) \quad (0.044) \quad (0.0023) \quad (0.0028) \quad (0.028)$$

$$n = 144 \quad R^2 = 0.93$$

The same model as from Multiple Regression Model 3 was used for the determination. It was interesting to see that the change is not very visible. Without those extreme values, the R-squared value dropped, which means that the explanation of the model actually got weaker. There were also no big differences of t-values and p-values for the significance. With this phenomenon, this paper assumes that this problem of increase of use of CO2 as a country's development is a world-wise problem.

#### C. Dummy Variables

As explained above, this paper added a dummy variable to figure out if there is any difference of results based on the developed or developing country basis. The standard use to divide the basis was the country's situation whether it is a member of OECD or not. The summary of the dummy variable is

provided on appendix A.

### Extra Multiple Regression Model 2

$$\log_{co2} = \beta_0 + \beta_1 \log\_gdppc + \beta_2 \text{renewable} + \beta_3 \text{accelec} + \beta_4 \log\_urbanpop + \beta_5 \text{oecd} + u$$

Equation 5:

$$\begin{aligned} \log_{co2} = & -8.56 + 0.30(\log\_gdppc) - 0.013(\text{renewable}) + 0.014(\text{accelec}) + 0.98(\log\_urbanpop) \\ & (0.78) \quad (0.059) \quad (0.0025) \quad (0.0029) \quad (0.029) \\ & + 0.053(\text{oecd}) \\ & (0.15) \end{aligned}$$

$$n = 148$$

$$R^2 = 0.94$$

The further illustration of the model is provided in appendix J. This regression model explains the further impact based on the situation whether the country is a member of OECD or not. As the number of oecd gets from 0 to 1, there is a 5.3% increase in the percentage change of CO2 emission. However, this interpretation must be done in careful sense since the p-value is extremely high with 0.753.

### VI. Conclusions

Throughout the result from the several regression models, this paper found that the primary *log\_gdp* variable remain positive coefficients without ruining the initial hypothesis that there will be a positive correlation between the country's CO2 emission and increase of GDP per capita. Even though it seems there can be some issues from the robustness test that may ruin the CLM assumptions, which can be critical, the model did explain quite high explanation with R-squared value almost reaching 0.94.

For the further implication from this paper, three things can be proposed. This paper only explains that the existence of correlation between a country's CO2 emission and economic development. Based on the result of this paper, one can look for further research of reasons of countries that are positioned higher than the regression line, finding for the reasons of those countries having higher CO2 emission than the average fit line.

In addition, as introduced in this paper, there are already a lot of research that are both done and on progress about the countries that are emitting high amount of CO2, best exemplified by 4 countries introduced in this paper – the US, China, Russia, and India. However, as shown in appendix H, there are several countries that are positioned in high GDP per capita, but remaining low emission of CO2, such as Luxembourg and Iceland. The benchmark research for those countries can be helpful to find for solutions a country can remain low CO2 emission with high economic growth.

Finally, since this paper only used the data from 2018 for the cross-country analysis there must be a limitation of understanding the change of trend as the time goes on. Therefore, time series analysis among the countries can be added, creating a harder extension of analysis of panel data of these countries. The analysis of panel data will provide the better understanding of the actual tendency of countries' CO<sub>2</sub> emission and economic development.

## Reference

- Douglas Holtz-Eakin, and Thomas M. Selden. "Stoking the fires? CO2 emissions and economic growth" *Journal of Public Economics*, vol 57, issue 1, 1995, pp 85-101.,  
[https://doi.org/10.1016/0047-2727\(94\)01449-X](https://doi.org/10.1016/0047-2727(94)01449-X)
- S.S. Wang, D.Q. Zhou, P. Zhou, and Q.W. Wang. "CO2 emissions, energy consumption and economic growth in China: A panel data analysis" *Energy Policy*, vol 39, issue 9, 2011, pp. 4870-4875.,  
<https://doi.org/10.1016/j.enpol.2011.06.032>
- Kais Saidi and Sami Hammami. "The impact of CO2 emissions and economic growth on energy consumption in 58 countries" *Energy Reports*, vol 1, 2015, pp 62-70.,  
<https://doi.org/10.1016/j.egyr.2015.01.003>.
- "World Development Indicators." *World Bank*, 2021,  
<https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators>
- "OECD member countries and partners." *OECD*, 2021  
<https://www.oecd.org/about/members-and-partners/>

## Appendix:

### Appendix A. Summary of Variables, STATA

```
summ logco2
```

Variable	Obs	Mean	Std. Dev.	Min	Max
logco2	148	10.07386	2.064254	5.736572	16.14896

```
summ log_gdppc
```

Variable	Obs	Mean	Std. Dev.	Min	Max
log_gdppc	148	8.769104	1.412213	6.165321	11.66648

```
summ renewable
```

Variable	Obs	Mean	Std. Dev.	Min	Max
renewable	148	32.75393	27.74344	0	96.3837

```
summ accelec
```

Variable	Obs	Mean	Std. Dev.	Min	Max
accelec	148	85.0282	25.32235	0	100

```
summ manuf
```

Variable	Obs	Mean	Std. Dev.	Min	Max
manuf	148	12.81661	6.472073	1.687851	39.91328

```
. summ log_urbanpop
```

Variable	Obs	Mean	Std. Dev.	Min	Max
log_urbanpop	148	15.65014	1.676542	12.07343	20.52947

summ oecd

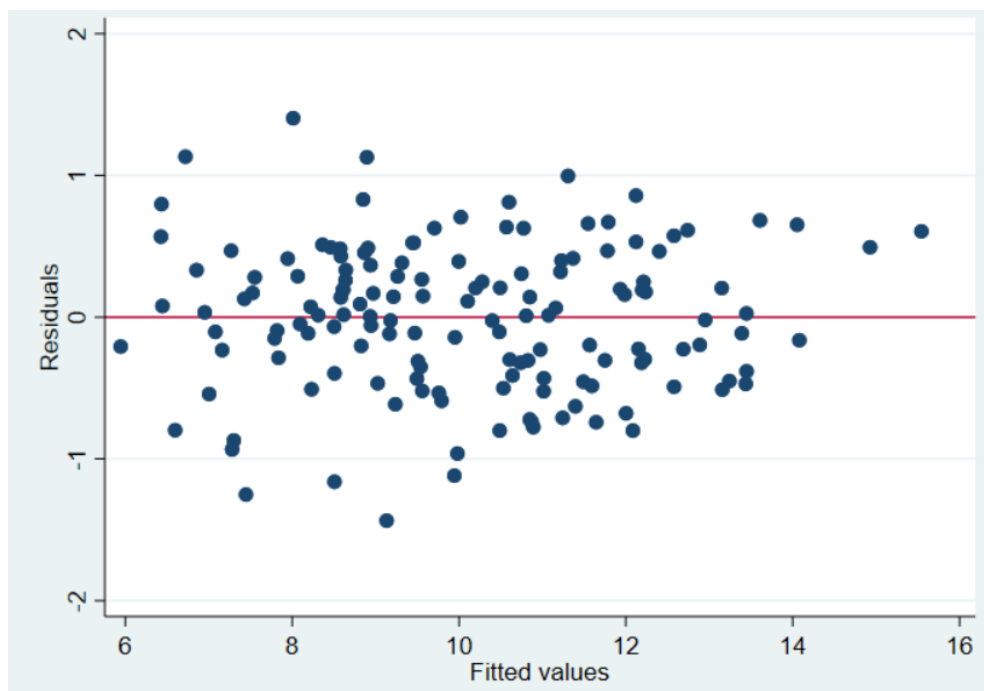
Variable	Obs	Mean	Std. Dev.	Min	Max
oecd	148	.2702703	.4456074	0	1

## Appendix B. Correlation test, STATA

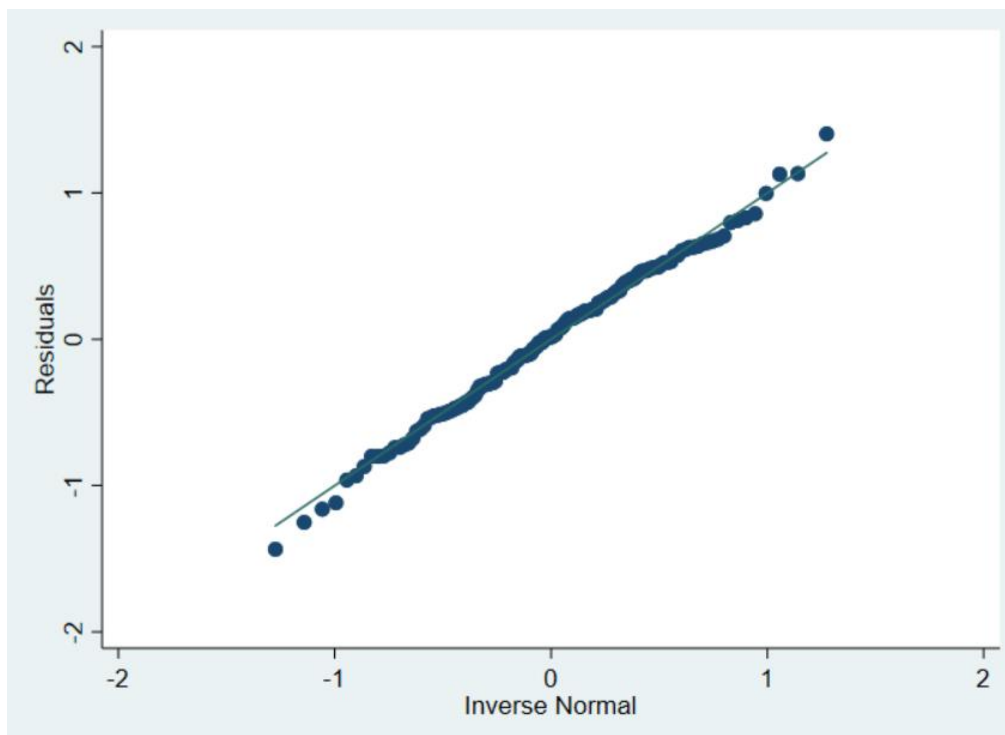
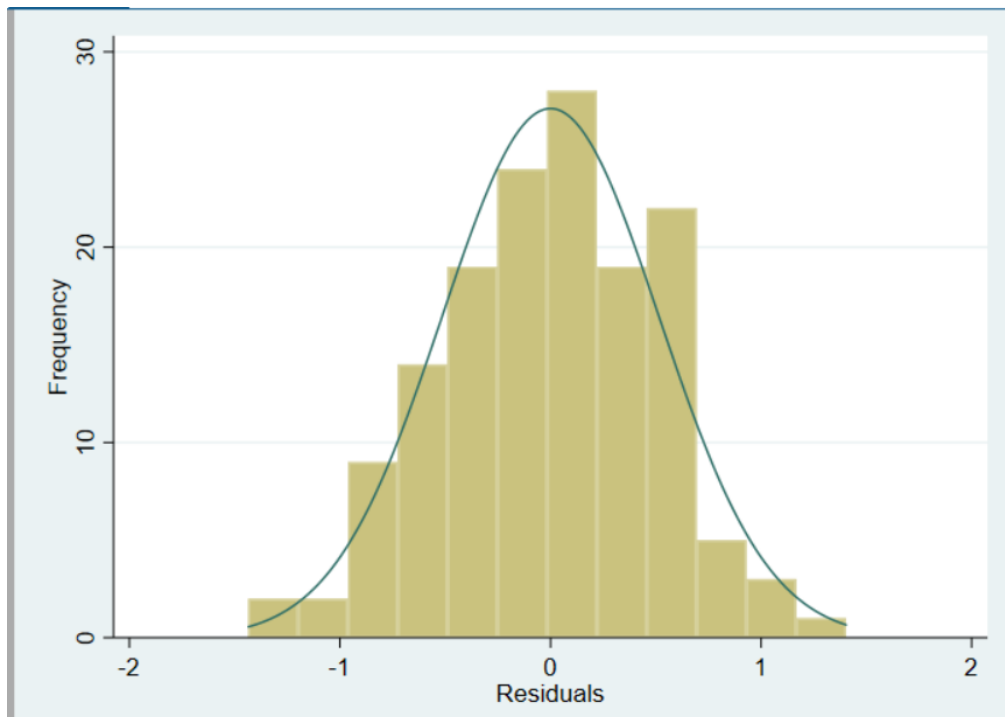
```
. corr log_gdppc renewable accelec manuf log_urbanpop  
(obs=148)
```

	log_gd~c	renewa~e	accelec	manuf	log_ur~p
log_gdppc	1.0000				
renewable	-0.5976	1.0000			
accelec	0.7011	-0.7360	1.0000		
manuf	0.1499	-0.1423	0.2242	1.0000	
log_urbanpop	0.0087	-0.1245	0.0720	0.2634	1.0000

## Appendix C. Residual vs fitted plot (y=0), STATA



## Appendix D. Histogram and Q-Q plot of residuals, STATA





## Appendix E. Regression Model 1, STATA

```
. regress logco2 log_gdppc
```

Source	SS	df	MS	Number of obs	=	148
Model	123.648035	1	123.648035	F(1, 146)	=	35.91
Residual	502.740166	146	3.4434258	Prob > F	=	0.0000
				R-squared	=	0.1974
				Adj R-squared	=	0.1919
Total	626.388201	147	4.26114422	Root MSE	=	1.8556

logco2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
log_gdppc	.6494337	.1083769	5.99	0.000	.4352436	.8636239
_cons	4.378911	.9625309	4.55	0.000	2.476617	6.281205

## Appendix F. Regression Model 2, STATA

```
. regress logco2 log_gdppc renewable accelec manuf log_urbanpop
```

Source	SS	df	MS	Number of obs	=	148
Model	586.581398	5	117.31628	F(5, 142)	=	418.49
Residual	39.8068025	142	.280329595	Prob > F	=	0.0000
				R-squared	=	0.9365
				Adj R-squared	=	0.9342
Total	626.388201	147	4.26114422	Root MSE	=	.52946

logco2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
log_gdppc	.3169329	.0442575	7.16	0.000	.2294442	.4044215
renewable	-.0138719	.0023777	-5.83	0.000	-.0185722	-.0091715
accelec	.0123838	.0030017	4.13	0.000	.0064501	.0183175
manuf	.0085263	.0071881	1.19	0.238	-.0056832	.0227359
log_urbanpop	.9713176	.027276	35.61	0.000	.917398	1.025237
_cons	-8.616073	.6027351	-14.29	0.000	-9.807566	-7.424579

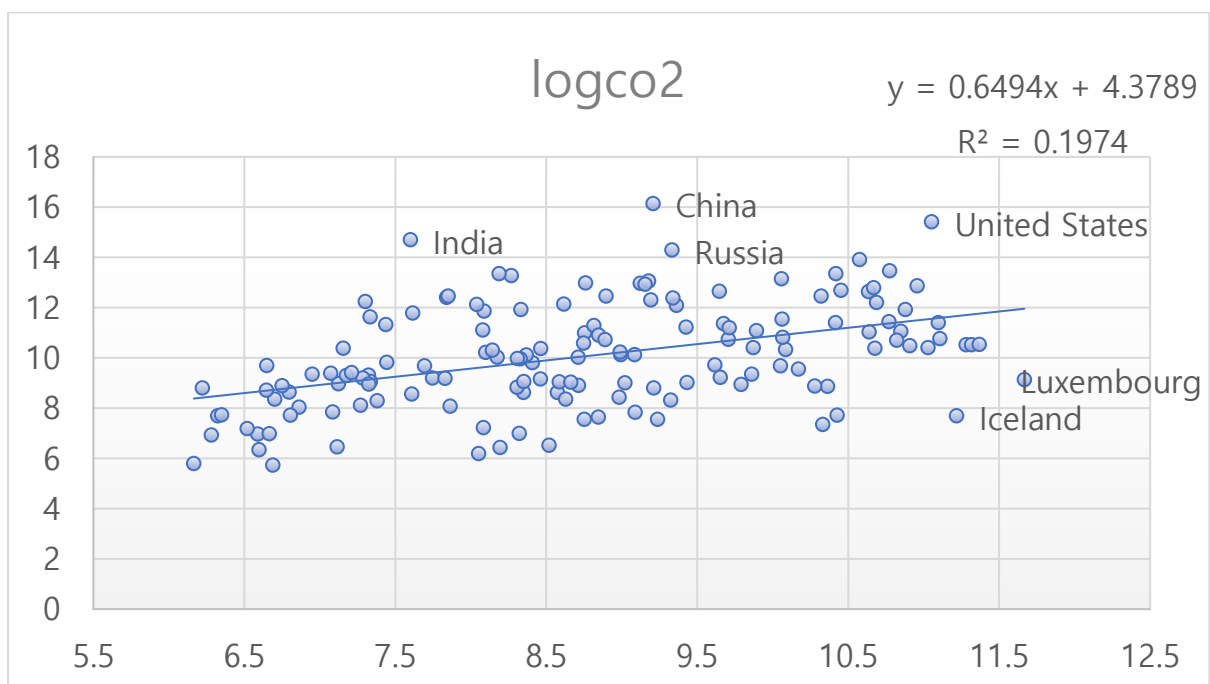
## Appendix G. Regression Model 3, STATA

```
. regress logco2 log_gdppc renewable accelec log_urbanpop
```

Source	SS	df	MS	Number of obs	=	148
Model	586.186973	4	146.546743	F(4, 143)	=	521.28
Residual	40.2012272	143	.281127463	Prob > F	=	0.0000
				R-squared	=	0.9358
				Adj R-squared	=	0.9340
Total	626.388201	147	4.26114422	Root MSE	=	.53021

logco2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
log_gdppc	.3174531	.0443182	7.16	0.000	.2298496	.4050567
renewable	-.0136703	.002375	-5.76	0.000	-.018365	-.0089756
accelec	.0129976	.0029609	4.39	0.000	.0071448	.0188505
log_urbanpop	.9796805	.0263866	37.13	0.000	.9275224	1.031839
_cons	-8.70111	.5993075	-14.52	0.000	-9.885756	-7.516463

## Appendix H. More precise scatter plot between CO2 emission and GDP per capita, Excel



## Appendix I. Extra Regression Model 1, STATA

```
. regress logco2 log_gdppc renewable accelec log_urbanpop
```

Source	SS	df	MS	Number of obs	=	144
				F(4, 139)	=	444.31
Model	481.173448	4	120.293362	Prob > F	=	0.0000
Residual	37.633065	139	.270741475	R-squared	=	0.9275
				Adj R-squared	=	0.9254
Total	518.806513	143	3.62801757	Root MSE	=	.52033

logco2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
log_gdppc	.3125736	.0441199	7.08	0.000	.2253408	.3998065
renewable	-.0131267	.0023408	-5.61	0.000	-.0177548	-.0084985
accelec	.0134262	.0028775	4.67	0.000	.0077369	.0191156
log_urbanpop	.9529082	.0281602	33.84	0.000	.8972304	1.008586
_cons	-8.311908	.6159734	-13.49	0.000	-9.529797	-7.094019

## Appendix J. Extra Regression Model 2, STATA

```
. regress logco2 log_gdppc renewable accelec log_urbanpop oecd
```

Source	SS	df	MS	Number of obs	=	148
				F(5, 142)	=	422.89
Model	586.969362	5	117.393872	Prob > F	=	0.0000
Residual	39.4188385	142	.277597454	R-squared	=	0.9371
				Adj R-squared	=	0.9349
Total	626.388201	147	4.26114422	Root MSE	=	.52688

logco2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
log_gdppc	.2984379	.0583389	5.12	0.000	.1831129	.413763
renewable	-.0134895	.0024503	-5.51	0.000	-.0183333	-.0086456
accelec	.0136831	.0028962	4.72	0.000	.0079579	.0194084
log_urbanpop	.9762711	.029209	33.42	0.000	.9185305	1.034012
oecd	.0526585	.145247	0.36	0.717	-.2344674	.3397843
_cons	-8.557801	.7824668	-10.94	0.000	-10.10459	-7.011012

#### **Appendix H. Countries included in the model (Alphabetic order)**

Albania, Algeria, Angola, Argentina, Armenia, Australia, Austria, Azerbaijan, The Bahamas, Bahrain, Bangladesh, Belarus, Belgium, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Burkina Faso, Cabo Verde, Cambodia, Cameroon, Central African Republic, Chad, Chile, China, Colombia, Congo Democratic Republic, Congo Republic, Costa Rica, Cote d'Ivoire, Croatia, Cuba, Cyprus, Czech Republic, Denmark, Djibouti, Dominican Republic, Ecuador, Egypt Arabian Republic, El Salvador, Equatorial Guinea, Estonia, Eswatini, Ethiopia, Fiji, Finland, France, Gabon, The Gambia, Georgia, Germany, Ghana, Greece, Guatemala, Guinea, Guinea-Bissau, Haiti, Honduras, Hungary, Iceland, India, Indonesia, Iran Islamic Republic, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea Republic, Kuwait, Kyrgyz Republic, Latvia, Lebanon, Lesotho, Liberia, Lithuania, Luxembourg, Malaysia, Maldives, Mali, Malta, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Montenegro, Morocco, Mozambique, Myanmar, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, North Macedonia, Norway, Oman, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russian Federation, Rwanda, Saudi Arabia, Senegal, Serbia, Sierra Leone, Singapore, Slovak Republic, Slovenia, South Africa, Spain, Sri Lanka, Suriname, Sweden, Switzerland, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Tunisia, Turkey, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uzbekistan, Vietnam, Zambia, Zimbabwe

#### **Appendix I. OECD + 5 key countries (based on year of 2018, alphabetic order)**

Australia, Austria, Belgium, Brazil, Chile, China, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Korea Republic, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, South Africa, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States